A new staging strategy for chronic obstructive pulmonary disease

Martin R Miller[1]
Ole F Pedersen[2]
Asger Dirksen[3]

[1]Department of Medicine, University Hospital Birmingham NHS Trust, Birmingham, UK; [2]Institute of Public Health, Department of Environmental and Occupational Medicine, University of Aarhus, Aarhus, DK; [3]Department of Pulmonary Medicine, Copenhagen University Hospital, Copenhagen, DK

Background: The best method for expressing lung function impairment is undecided. We tested in a population of patients with chronic obstructive pulmonary disease (COPD) whether forced expiratory volume in 1 second (FEV_1) or FEV_1 divided by height squared (FEV_1/ht^2) was better than FEV_1 percent predicted (FEV_1PP) for predicting survival.

Method: FEV_1, FEV_1PP, and FEV_1/ht^2 recorded post bronchodilator were compared as predictors of survival in 1095 COPD patients followed for 15 years. A staging system for severity of COPD was defined from FEV_1/ht^2 and compared with the Global Initiative for Obstructive Lung Disease (GOLD) staging system.

Result: FEV_1/ht^2 was a better univariate predictor of survival in COPD than FEV_1 and both were better than FEV_1PP. The best multivariate model for predicting survival included FEV_1/ht^2, age and sex. Comparing the GOLD stages with the FEV_1/ht^2 groups found that survival was more coherent within each FEV_1/ht^2 group than it was within each GOLD stage. FEV_1/ht^2 had 60% more people in its most severe group than the severest GOLD stage with these extra subjects having equivalently poor survival and had 155% more in the least severe group with equivalent survival. GOLD staging misclassified 51% of subjects with regard to survival.

Conclusion: We conclude that GOLD criteria using FEV_1PP do not optimally stage COPD with regard to survival. An alternative strategy using FEV_1/ht^2 improves the staging of this disease. Studies which stratify COPD patients to determine the effect of interventions such as drug trials, rehabilitation, or management guidelines should consider alternatives to the GOLD classification.

Keywords: chronic obstructive pulmonary disease, spirometry, respiratory function tests

Introduction

Lung function measurements are made to help determine if a subject has an abnormality of lung function and also to monitor a disease or its response to treatment. The degree of any abnormality may be needed to determine choices about therapy or level of disability benefit but may also be needed to estimate prognosis. To determine if a lung function result is abnormal the American Thoracic Society (ATS) and European Respiratory Society (ERS) recommend (Quanjer et al 1993; ATS 1994) that the individual's lung function data should be compared with a predicted value using the method of standardized residuals (SR). SR are calculated from:

$$(Recorded\ FEV_1 - Predicted\ FEV_1)/RSD$$

where RSD is the residual standard deviation from the regression equation used for the prediction, so that an FEV_1SR of -1.645 is at the lower 90% confidence limit of normality.

When it comes to expressing the degree of any such abnormality, that is the severity of impairment, there is no agreed method that gives equivalent results for all ages and both sexes. The method of percent of predicted (PP) is widely used for expressing disease severity (ATS 1987) but its validity is not based on statistical evidence (Sobol and Weinheimer 1966; Miller and Pincock 1988). SR are potentially

Correspondence: Martin R Miller
Department of Medicine, University of Birmingham, Selly Oak Hospital, Birmingham B29 6JD, UK
Tel +44 121 627 8479
Fax +44 121 627 8292
Email martin.miller@uhb.nhs.uk

flawed for expressing severity in that a young subject's predicted value is much higher than that for an older subject, so a young subject can have a forced expiratory volume in 1 second (FEV_1) 5 SR below predicted which for an older subject would never be found as the FEV_1 would have to be below zero.

The PP method has been used in many epidemiological studies (Lange et al 1990; Knuiman et al 1999; Schunemann et al 2000; Thomason and Strachan 2000) showing that FEV_1 expressed as percent predicted (FEV_1PP) is an important predictor of survival. However, in chronic obstructive pulmonary disease (COPD) it was raw FEV_1 that was originally shown to relate to survival (Burrows and Earle 1969). The Framingham study found that FEV_1 divided by height squared was an important predictor of survival (Ashley et al 1975) in the general population, a finding recently reinforced from the Reykjavik study (Chinn et al 2007). More recently the PP method has been used with arbitrary threshold values in guidelines with regard to specifying the severity of COPD (ATS 1995; Siafakas et al 1995) and asthma (British guideline 2005) and the Global Initiative for Chronic Obstructive Lung disease (GOLD) has empirically defined 5 stages of severity (NIH et al 2005) based on FEV_1PP. Concern has been raised that this system is not evidence-based (Kerstjens 2004).

Prediction equations such as those recommended by the ERS (Quanjer et al 1993) are used to derive a predicted value for an individual which tries to take into account the effect of age, height, and sex on lung function, but this is imprecise with wide confidence intervals and the data from which the equations are derived may not perfectly match the patient under consideration. Therefore any method of relating lung function to a predicted value may introduce some age, height or sex related error. We have therefore tested whether lung function impairment expressed as absolute FEV_1 or standardized by height might be better than FEV_1PP and FEV_1SR in predicting survival in COPD.

Methods

We have reanalyzed the data from 1095 subjects with COPD who had been recruited from 1983 to 1988 in a respiratory clinic in Copenhagen into a longitudinal study to look at predictors of their all cause mortality (Hansen et al 1999, 2001) over a 15 year follow up. They were all assessed by a consultant respiratory physician and diagnosed as having COPD based from symptoms of cough, sputum, lack of variability in airflow limitation, lack of allergy and atopy, their smoking history and their FEV_1 as a percent of forced

Table I Classification criteria for GOLD and FEV_1/ht^2 staging from spirometry alone. $FEV_1\%$ is the FEV_1 expressed as a percentage of FVC. FEV_1PP is the FEV_1 expressed as a percentage of the subject's predicted value

GOLD staging		FEV_1/ht^2 staging	
Stage	Criteria	Stage	Criteria
0	$FEV_1\% \geq 70\%$		
I	$FEV_1\% < 70\%$ and $FEV_1PP \geq 80\%$	I	$0.5 < FEV_1/ht^2$
2	$FEV_1\% < 70\%$ and $FEV_1 PP < 80\%$	2	$0.4 < FEV_1/ht^2 \leq 0.5$
3	$FEV_1\% < 70\%$ and $FEV_1PP < 50\%$	3	$0.3 < FEV_1/ht^2 \leq 0.4$
4	$FEV_1\% < 70\%$ and $FEV_1PP < 30\%$	4	$FEV_1/ht^2 \leq 0.3$

vital capacity (FVC) had to be less than 89% of their predicted value, in accordance with ERS recommendations at that time (Siafakas et al 1995). There was no within day variation in lung function and no significant acute or corticosteroid reversibility to their airflow obstruction. We have analyzed the spirometric data on entry into the study which were obtained after both an inhaled bronchodilator and a 2 week trial of oral corticosteroids, ie, maximally bronchodilated as per GOLD recommendation (NIH et al 2005).

Each subject's data had their FEV_1 expressed as percent predicted (PP) using ERS prediction equations (Quanjer et al 1993), as FEV_1 divided by height squared (FEV_1/ht^2) and as standardized residuals (SR) (Miller et al 1985; Quanjer et al 1993). Subjects were staged by GOLD criteria (NIH et al 2005) and also arbitrarily by FEV_1/ht^2 into 4 stages as defined in Table 1. Because GOLD stage 1 only identified 6 subjects in this cohort the staging by FEV_1/ht^2 was into 4 and not 5 arbitrary groups in order to allow fairer comparison between the two classification systems.

Statistical analysis was undertaken using SPSS for Windows version 11.5 (Microsoft, Redmond, WA). Cox proportional hazards regression analysis was used for predicting all cause mortality using quintiles of putative predictor variables. This method is a special form of multiple regression that allows for the inclusion of the censored data for those cases for which the end point of death has not yet occurred. The best prediction models were judged on the basis of highest global chi-square (χ^2) value and greatest reduction in –2 times the log likelihood (–2LL) for different models (Cox 1972). The proportional hazards model assumes that the ratio of the estimated hazard functions for any two observations with different values for the independent (predictor) variables is constant over time. This assumption was checked for each

Table 2 Descriptive data for age, height and spirometric values for the 1095 patients in the study (644 female) showing mean, standard deviation (SD) and median values

	Mean (SD)	Median
Age	61.8 (9.9)	63.0
Ht in meters	1.66 (0.09)	1.65
FVC in L	2.25 (0.76)	2.14
FVCSR	−1.87 (1.10)	−1.86
FEV_1 in L	1.21 (0.45)	1.14
FEVSR	−3.21 (1.02)	−3.24
FEV_1PP	46.7 (14.5)	45.5
FEV_1/ht^2	0.43 (0.15)	0.41
$FEV_1\%$	54.4 (12.2)	55.0
$FEV_1\%SR$	−3.31 (1.76)	−3.23
PEF in L/s	3.60 (1.47)	3.33
PEFSR	−3.11 (1.32)	−3.28

model from plots of cumulative hazard against survival time stratified by the relevant FEV_1 index. For the multivariate Cox models, hazard ratios were calculated by contrasting risk against that for the first stage of each index. Subjects were categorized by their GOLD and FEV_1/ht^2 stages and Kaplan-Meier survival analysis was used to estimate mean survival for every subgroup of subjects defined by the 4×4 matrix of GOLD and FEV_1/ht^2 stages. Mean survival was used rather than median survival since some subgroups were too small for reliable estimates of the median.

Results

Age, height, and spirometric values are shown in Table 2 for the 1095 subjects of whom 644 (59%) were female. The median survival from entry into the longitudinal study was 3208 days, with 723 (66%) dying during the period of follow up (median survival 2231 days versus 4083 days for survivors), of whom 404 were female which was slightly lower than the number (425) expected by chance ($\chi^2 = 7.57$, df = 1, p = 0.006). Univariate Cox proportional hazards models for survival were derived with the following variables as predictors: height, sex, body mass index (BMI), FEV_1, FEV_1/ht^2, FEV_1PP and FEV_1SR. Table 3 shows the strength of univariate association with survival for each index as a continuous variable with FEV_1/ht^2 being the best univariate predictor.

To allow hazard ratios for various predictors to be calculated Cox regression models were constructed with sex and quintiles of age and the various lung function indices. The best multivariate model was with FEV_1/ht^2, age and sex ($\chi^2 = 289$, −2LL = 9025) and the hazard ratios for the predictors are shown in Table 4. The other FEV_1 indices

gave less good models with the next best being with raw FEV_1, followed by FEV_1PP and then FEV_1SR. The model with FEV_1/ht^2 had a lower HR for male sex than that with raw FEV_1 suggesting the standardization by height had taken into account some sex difference. The models with PP and SR found the HR for male sex was no longer significant. BMI quintiles improved each of these models slightly but only the lightest BMI quintile had a significantly increased hazard ratio (HR = 1.4, 95% CL 1.1 to 1.8) compared with the other BMI quintiles.

Figure 1 shows mean survival (with 95% CL bars) for each FEV_1/ht^2 group compared with the GOLD stage. Figure 2 shows the numbers of subjects in each of the staging groups and the full data used in these Figures is presented in Table 5. The mean survival for the FEV_1/ht^2 groups decreases progressively and fairly evenly whereas the survival for GOLD 0 and 2 are very similar. There were 308 subjects in FEV_1/ht^2 stage 1 which was 2.5 times the number of subjects in the equivalent GOLD stage 0 (121). More than half of the subjects in each of GOLD stages 2 and 3 were in discordant FEV_1/ht^2 stages and the worst FEV_1/ht^2 stage included 1.6 times as many subjects as did GOLD stage 4 (199 versus 126). Overall 51% of subjects were in discordant staging groups. The poor staging by GOLD was confirmed from a Cox regression model using just the GOLD stage as predictor for survival, where the χ^2 value was only 81, which was significantly inferior to a model using just the FEV_1/ht^2 stage which yielded χ^2 value of 166.

Discussion

We present the first evidence that FEV_1PP is not as good as either raw FEV_1 or FEV_1/ht^2 at predicting survival in a cohort of COPD patients. This finding confirms concerns previously expressed about the use of PP as a method for expressing

Table 3 *Univariate analysis*: Indices that were found by Cox regression to be univariate predictors of survival with χ^2 value for strength of prediction and its significance value

Index	χ^2 value	p value
FEV_1/ht^2	161	<0.0001
Age	159	<0.0001
FEV_1	127	<0.0001
FEV_1PP	111	<0.0001
$FEV_1\%$	97	<0.0001
FEV_1SR	18	<0.0001
Sex	12	<0.001
BMI	11	<0.001
Weight	8	0.004
Height	0	

Table 4 *Multivariate analysis:* The table shows the covariate hazard ratios (HR) for death (with 95% confidence limits) for Cox proportional hazards models using sex and the quintiles of age and various methods for expressing FEV_1 impairment. The HR values are contrasts against the hazard estimated for the first stage of each group and for male contrasted with female. Chi-square values are shown for each model

|		FEV_1/ht^2 Chi square = 289		Raw FEV_1 Chi square = 276		FEV_1PP Chi square = 269		FEV_1SR Chi square = 218	
		HR	95% CL	HR	95% CL	HR	95% CL	HR	95% CL
Age	youngest quintile	1.0		1.0		1.0		1.0	
	quintile 2	1.8	(1.3 to 2.4)	1.8	(1.3 to 2.4)	2.0	(1.5 to 2.6)	2.1	(1.6 to 2.9)
	quintile 3	2.2	(1.6 to 2.9)	2.2	(1.7 to 3.0)	2.6	(2.0 to 3.5)	3.1	(2.4 to 4.2)
	quintile 4	2.5	(1.8 to 3.3)	2.5	(1.9 to 3.3)	3.2	(2.4 to 4.3)	4.1	(3.1 to 5.4)
	oldest quintile	3.1	(2.4 to 4.1)	3.1	(2.3 to 4.1)	4.3	(3.3 to 5.7)	6.0	(4.5 to 7.9)
FEV_1 index	best quintile	1.0		1.0		1.0		1.0	
	quintile 2	1.6	(1.2 to 2.1)	1.6	(1.2 to 2.1)	1.5	(1.1 to 1.9)	1.2	(0.9 to 1.5)
	quintile 3	2.0	(1.5 to 2.7)	1.9	(1.4 to 2.5)	1.8	(1.4 to 2.4)	1.6	(1.3 to 2.0)
	quintile 4	2.4	(1.8 to 3.2)	2.4	(1.8 to 3.2)	2.4	(1.8 to 3.0)	2.0	(1.5 to 2.5)
	worst quintile	4.0	(3.0 to 5.2)	3.9	(3.0 to 5.1)	3.4	(2.6 to 4.3)	2.6	(2.0 to 3.3)
Male		1.5	(1.3 to 1.7)	1.8	(1.5 to 2.0)	1.1	(0.9 to 1.3)	1.1	(0.9 to 1.3)
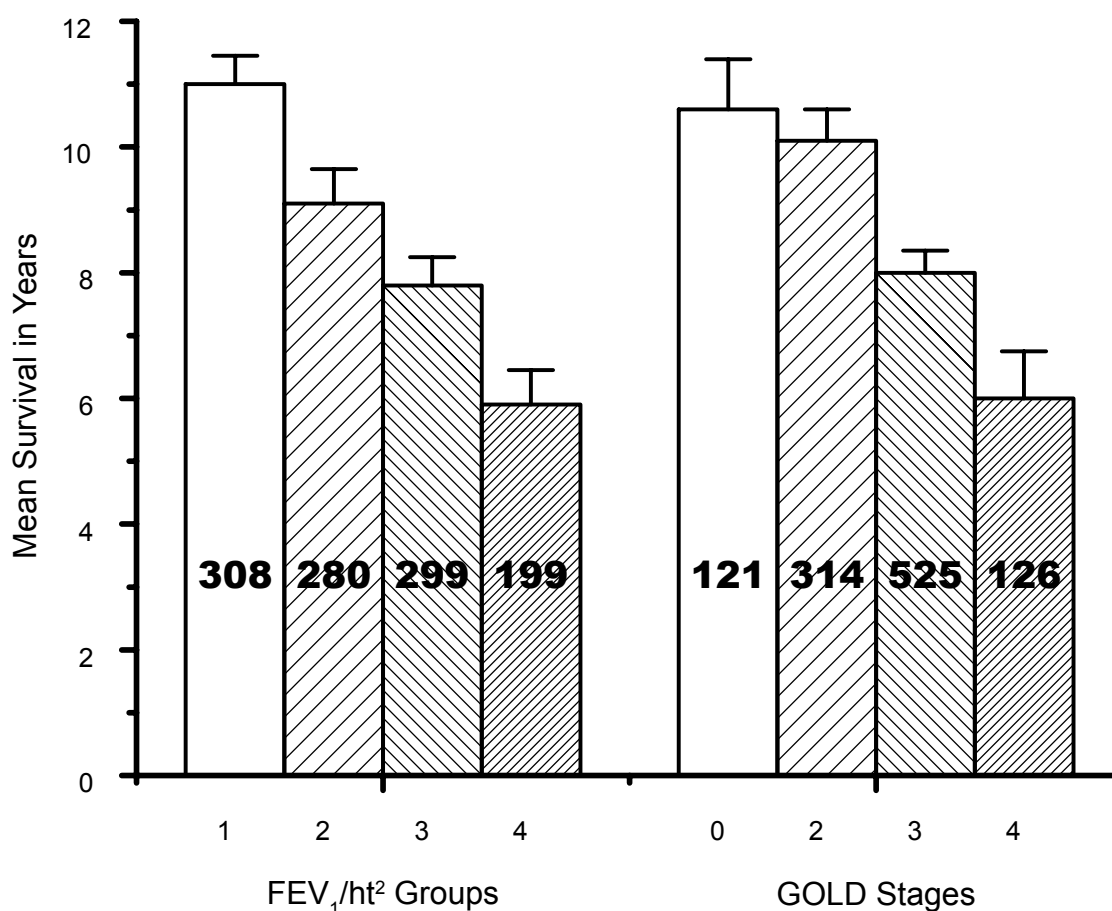
Figure 1 The mean survival for each FEV_1/ht^2 group and each GOLD stage with 95% confidence limit bars and the number of subjects in each column.
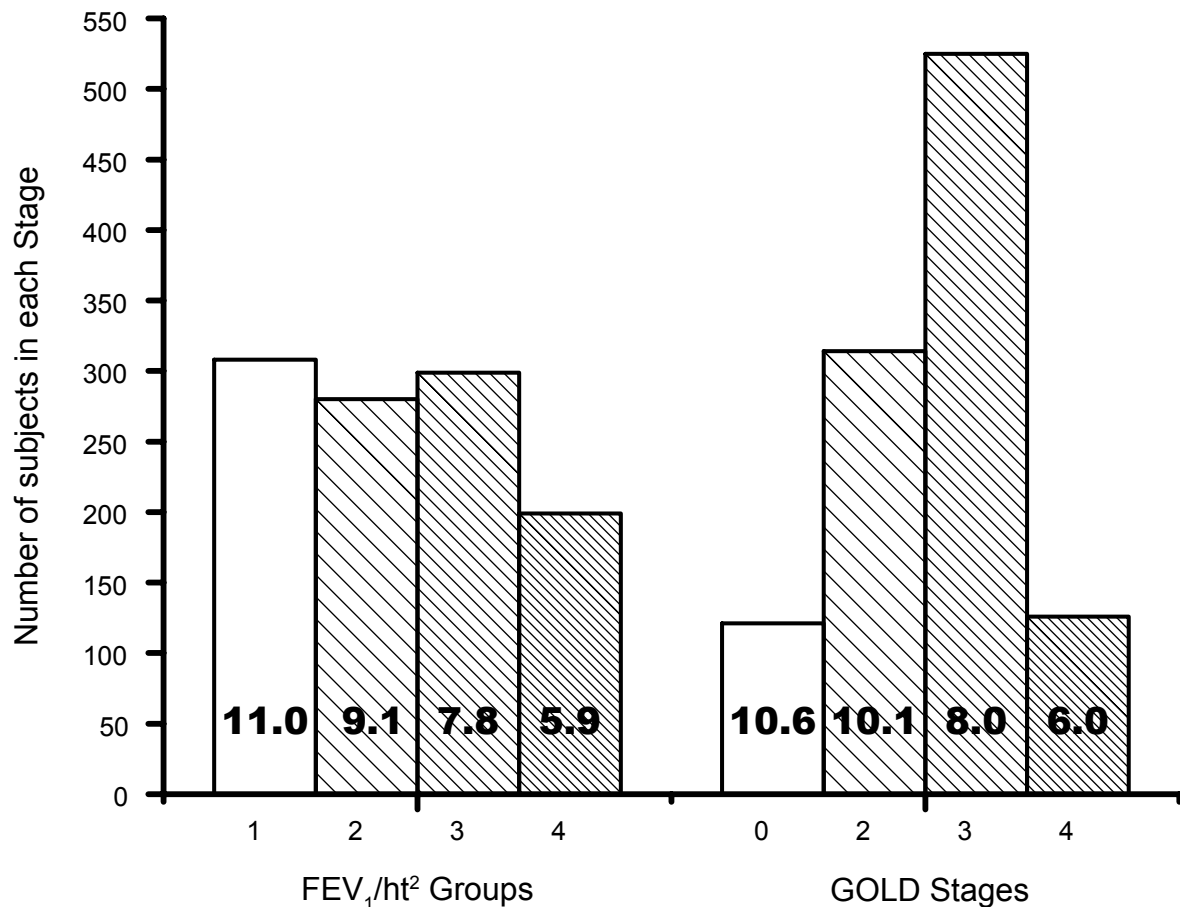
Figure 2 Plot of the number of subjects in each of the groups for the FEV_1/ht^2 staging system on the left and the GOLD staging system on the right with the mean survival of each group stated inside the relevant column.

impairment in FEV_1 (Sobol and Weinheimer 1966; Miller and Pincock 1988). The assumption that a fixed level of FEV_1PP means the same level of impairment for different subjects is clearly flawed since, for example, young adults with cystic fibrosis can survive with an FEV_1 just as low in absolute terms as old subjects, with their PP values going as low as 10% (Sood et al 2001) which is much lower than that seen in the elderly.

Several studies have shown that survival in COPD relates to FEV_1 (Burrows and Earle 1969; Gorecka et al 1997; Thomason and Strachan 2000) with values lower than 0.75 litre having a 3 year survival of only 50% (Burrows and Earle 1969). When using the GOLD criteria to stratify our subjects there were only 11.5% subjects in stage 4 whereas the method based on FEV_1/ht^2 increased the proportion in the severest stage to 18.2% and yet these had a comparably poor survival. This suggests that estimating severity of COPD from FEV_1PP criteria does not identify all those subjects who are most severely affected. Concern has been expressed that the GOLD criteria, which were arbitrarily

chosen to define severity of COPD, are not based on any evidence of their ability to predict survival or any other aspect of COPD management (Kerstjens 2004). Our data provide strong evidence that the GOLD criteria are not adequate for correctly placing individual patients into disease stage categories that relate to survival. We do not have data on individual symptoms to determine how the two classification systems of COPD relate to symptoms and performance but this could be verified in other data sets.

Furthermore we found that within the GOLD stages there were subgroups with differing survival that were better stratified by FEV_1/ht^2 staging. The definition of stage 4 by GOLD criteria severity (NIH et al 2005) can include subjects with less severe spirometry results but who are in type 2 respiratory failure. Therefore it is possible that some of our subjects allocated to GOLD 3 might in fact be in GOLD stage 4 if their arterial gas results were known. However, it seems unlikely that the 85 subjects in the worst FEV_1/ht^2 stage but who we classified as GOLD 3 had been in type 2 respiratory failure on entry into the study because their

Table 5 Cross tabulation of concordance in the 1095 subjects between GOLD classification (GOLD stages 0 to 4) and FEV$_1$/ht^2 classification for COPD severity. The table shows the number of subjects dying (nd) under follow up as a fraction of the total number N within that group (nd/N), the estimated Kaplan-Meier mean survival for that group as years in bold and the 95% confidence limits for the estimate in parenthesis. Aggregate results for each columns and row as a whole are also shown

| | Staging by FEV$_1$/ht^2 | | | | | | | | | | | | | | | |
	Stage 1 (>0.5)			Stage 2 (>0.4 and ≤0.5)			Stage 3 (>0.3 and ≤0.4)			Stage 4 (≤0.3)					
GOLD 0	29/75	**11.2**	(10.2–12.2)	13/29	**9.8**	(8.4–11.3)	12/15	**9.0**	(7.1–10.8)	2/2	**5.7**	(0.0–13.3)	56/121	**10.6**	(9.8–11.4)
GOLD 2	88/195	**10.8**	(10.2–11.4)	66/101	**9.3**	(8.5–10.2)	15/18	**6.2**	(4.7–7.6)	–	–	–	169/314	**10.1**	(9.6–10.6)
GOLD 3	17/37	**10.7**	(9.6–11.9)	101/150	**8.7**	(8.0–9.4)	194/253	**7.9**	(7.4–8.4)	76/85	**6.1**	(5.2–6.9)	388/525	**8.0**	(7.7–8.4)
GOLD 4	1/1	–	–	–	–	–	11/13	**6.6**	(4.3–8.9)	98/112	**5.8**	(5.1–6.5)	110/126	**6.0**	(5.2–6.7)
	135/308	**11.0**	(10.5–11.4)	180/280	**9.1**	(8.6–9.7)	232/299	**7.8**	(7.4–8.3)	176/199	**5.9**	(5.4–6.5)			

mean survival was 6.1 years. It seems clear that FEV$_1$/ht^2 can stratify subjects with COPD into survival groups better than GOLD criteria without recourse to other data such as arterial gases.

Recent work has emphasized that multivariate analysis including data on BMI, the degree of airflow obstruction as FEV$_1$PP, dyspnoea, and exercise capacity measured by the six-minute-walk test (BODE index) was better at predicting survival in COPD than univariate analysis (Celli et al 2004). However, our data suggest that the use of FEV$_1$/ht^2 instead of FEV$_1$PP as the index of airflow obstruction in this form of multivariate analysis would further improve the survival prediction.

Additional problems with the PP and SR methods for expressing degree of abnormality occur because they require an estimate of a subject's predicted value. The regression equations recommended by the ERS for FEV$_1$ only account for about 58% of the variation of FEV$_1$ in normal subjects, thus the predicted value is not a precise estimate and will include errors related to sex, age, height, and technical issues (Quanjer et al 1993). Furthermore if the subject whose predicted value is being derived is from a population different from that used to derive the prediction equation or the subject's age is outside the limits of the population used in the equation then further errors will be incurred.

Our data show FEV$_1$/ht^2 is better than both raw FEV$_1$ and FEV$_1$PP for expressing degree of lung function impairment. If this dataset is analyzed with the two sexes considered independently then FEV$_1$/ht^2 gives results indistinguishable from that for raw FEV$_1$ in terms of ability to predict all cause mortality. However, in clinical practice it is advantageous to have a method for assessing risk that is applicable for both sexes in an equivalent way. FEV$_1$/ht^2 is one way of taking some of the sex differences in lung function into account without introducing the potential errors inherent in trying to predict what a given subject's FEV$_1$ should be (Quanjer et al 1993). A recent study of all cause mortality in a general population confirmed that FEV$_1$/ht^2 was better related to survival than FEV$_1$PP (Chinn et al 2007) and we now confirm this finding in a cohort of COPD patients.

We conclude that FEV$_1$/ht^2 is better related to survival in COPD than FEV$_1$PP and so may be the best method for expressing degree of FEV$_1$ impairment. Our findings indicate that applying the GOLD criteria to COPD management does not optimally classify those most severely affected. So studies which need to stratify COPD patients in order to determine the potential benefits of interventions such as drug trials, rehabilitation programmes or even discharge policies should

not just consider the GOLD classification but also consider using FEV_1/ht^2 or other strategies to assess disease severity correctly. Our evidence suggests using GOLD criteria alone will have suboptimal power to show any benefits from such interventions. We believe future scientific endeavor in COPD must not be limited by the arbitrary GOLD staging which here failed to identify over one third of the worst prognosis group. Alternative classifications should now be prospectively compared so that the best results for managing and researching into COPD can be achieved.

References

[ATS] American Thoracic Society. 1987. Standardization of spirometry: 1987 update. *Am Rev Respir Dis*, 136:1286–96.

[ATS] American Thoracic Society. 1995. Standardization of spirometry. 1994 update. *Am J Respir Crit Care Med*, 152:1107–36.

[ATS] American Thoracic Society. 1995. COPD Guidelines. *Am J Respir Crit Care Med*, 152:S77–S120.

Ashley F, Kannel WB, Sorlie PD, et al. 1975. Pulmonary function: relation to aging, cigarette habit, and mortality. The Framingham Study. *Ann Intern Med*, 82:739–45.

British Guideline. 2003. British guideline on the management of asthma. *Thorax*, 58:Suppl 1.

Burrows B, Earle RH. 1969. Course and prognosis of chronic obstructive lung disease. A prospective study of 200 patients. *N Engl J Med*, 280:397–404.

Burrows B, Earle RH. 1969. Prediction of survival in patients with chronic airway obstruction. *Am Rev Respir Dis*, 99:865–71.

Celli BR, Cote CG, Marin JM, et al. 2004. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. *N Engl J Med*, 350:1005–12.

Chinn S, Gislason C, Aspelund T, et al. 2007. Optimum expression of adult lung function based on all-cause mortality: results from the Reykjavik study. *Respir Med*, 101:601–9.

Cox DR. 1972. Regression models and life tables. *J Royal Stat Soc*, B34:187–220.

Gorecka D, Gorzelak K, Sliwinski P, et al. 1997. Effect of long term oxygen therapy on survival in patients with chronic obstructive pulmonary disease with moderate hypoxaemia. *Thorax*, 52:667–8.

Hansen EF, Phanareth K, Laursen LC, et al. 1999. Reversible and irreversible airflow obstruction as predictor of overall mortality in asthma and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, 159:1267–71.

Hansen EF, Vestbo J, Phanareth K, et al. 2001. Peak flow as predictor of overall mortality in asthma and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, 163:690–3.

Kerstjens HAM. 2004. The GOLD classification has not advanced understanding of COPD. *Am J Respir Crit Care Med*, 170:212–13.

Knuiman MW, James AL, Divitini ML, et al. 1999. Lung function, respiratory symptoms, and mortality: results from the Busselton Health Study. *Ann Epidemiol*, 9:297–306.

Lange P, Nyboe J, Appleyard M, et al. 1990. Spirometric findings and mortality in never-smokers. *J Clin Epidemiol*, 43:867–73.

Miller MR, Pincock AC, Grove DM. 1985. Patterns of spirogram abnormality in individual smokers. *Am Rev Respir Dis*, 132:1034–40.

Miller MR, Pincock AC. 1988. Predicted values: how should we use them? *Thorax*, 43:265–7.

[NIH] National Institutes of Health, National Heart, Lung and Blood Institute. 2005. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Workshop Report. Update 2005 [online]. Accessed on May 22, 2007. URL: http://www.goldcopd.com

Quanjer PhH, Tammeling GJ, Cotes JE, et al. 1993. Standardized lung function testing. Lung volumes and forced ventilatory flows. *Eur Respir J*, 6, (Suppl 16):5–40.

Schunemann HJ, Dorn J, Grant BJ, et al. 2000. Pulmonary function is a long-term predictor of mortality in the general population: 29-year follow-up of the Buffalo Health Study. *Chest*, 118:656–64.

Siafakas NM, Vermeire P, Pride NB, et al. 1995. ERS – Consensus statement. Optimal assessment and management of chronic obstructive pulmonary disease (COPD). *Eur Respir J*, 8:1398–420.

Sobol BJ, Weinheimer A. 1966. Assessment of ventilatory abnormality in the asymptomatic subject: an exercise in futility. *Thorax*, 21:445–9.

Sood N, Paradowski LJ, Yankaskas JR. 2001. Outcomes of intensive care unit care in adults with cystic fibrosis. *Am J Respir Crit Care Med*, 163:335–8.

Thomason MJ, Strachan DP. 2000. Which spirometric indices best predict subsequent death from chronic obstructive pulmonary disease? *Thorax*, 55:785–8.